Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Graphic Language Documents Structures and Functions

Caroline Beebe
North Carolina State University
Raleigh, North Carolina

Elin Jacob
School of Library and Information Science
Indiana University
Bloomington, Indiana

Abstract

This paper explores the nature of graphic language documents from the perspective of the document's structure as a spatially-oriented object in order to identify existing and potential methodologies for representation and classification of images in the electronic environment. Drawing upon the Bauhaus concept that form (or structure) follows function and design principles derived from Gestalt theory, it addresses the relationship that exists between structure and function in the broad domain of graphic language documents.

Introduction

> The human preference for image over text, dating back at least 40,000 years to our Stone Age ancestors, continues to inspire the mysteries and technology of human communications and information storage. Over time, we've learned that written words are easier to store and disseminate, ultimately propelling text to the status of medium of choice. But technological developments are putting images back in the picture as new capabilities allow for creating, storing, managing, and distributing visual information in images and video. (Diane Crawford)

There is general agreement that the term *document* must be understood to include not only written text but graphic images, pictures, and objects as well. Twyman's (1979) definition of graphic language as two-dimensional, intentional, visible communication leads to an understanding of graphic language documents as subsuming any traditional alphanumeric text as well as any combination of text, graphics, and/or pictures. In the digital environment, Twyman's definition of graphic language documents can be applied to all displays of signs and symbols to reconcile the traditional text-based document with the image as document. Within a digital system, there is no difference between the representation of an alphanumeric text and a picture: each format simply requires a different internal code for processing the ones and zeros that constitute the digital display. Each letter code is individually manifested as an array of pixels or picture elements that conform to Gestalt principles of structural relationship to create the image on the digital display: that is, each pixel is either on or off so that the spatial arrangement of the one/off picture

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

elements creates meaningful forms. All standard digital displays, regardless of specific content, are actually graphic language documents and are therefore subsumed within the concept of image as defined in this paper.

In electronic environments, there are increasing numbers of resources that are not strictly alphanumeric. The dominance of text-based documents is being challenged as the availability and use of non-text, or more-than-text, documents expands in digital environments such as the World Wide Web. Within this evolving context, methods for storing and retrieving graphic language documents in electronic systems are based on the common assumption that they can be represented only through words -- through text-based, linguistic representations. Consideration must be given both to the adequacy of linguistic representation as the predominant means of organizing and accessing graphic language documents and to the potential for developing additional approaches to representation that can apply to all forms of documents.

Traditional representational structures strive to achieve consistency through stability and in so doing impose an internal rigidity on information storage and retrieval systems. When applied to text documents, physical description (cataloging) has little impact on information gleaned from contextual and conceptual interpretations of the document's content. Within an expanded understanding of graphic language documents, physical structure need not be independent of the meaningful content of a document. When it is assumed that the specific arrangement of objects and entities within a spatial or pictorial process of communication is equivalent to the arrangement of words that form a sentence or a paragraph in a process of verbal communication, the physical arrangement of entities within such a framework holds potential meaning that cannot be captured in standard verbal representations.

Within the context of the electronic environment, Chang (1986) identifies four categories of visual languages based on the structures required by a computer system. These categories are described by their purpose:

1) for programming with visual expressions
   (logical objects with visual representations)
2) for supporting visual interactions
   (logical objects with visual representation)
3) for processing visual information
   (visual objects with imposed logical representation)
4) for iconic visual information processing
   (visual objects with imposed logical representation)

The first two categories have applications in image processing, computer vision, robotics, image database management, image communications, etc. The third and fourth categories have applications in computer graphics, interface design, form management, and computer aided design. This approach to visual languages is based on the artificial reality of the mechanics of the computer system: these languages exist because they have been created within the structure of the computer system. They are dependent on an imposed logical representation; and that representation must be translated from the real world of the object, which is itself a surrogate, to the logic of the computer system. This is the traditional role of the cataloguer and the indexer in

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

information storage and retrieval systems. With the development of image processing and computer vision capabilities, some of this activity can be handled automatically by the system itself. But such automation of representation must be informed by its relationship to the real world. This paper attempts to build upon this relationship by suggesting categories of physical structure that can be identified by the computer system and subsequently used for representing non-text-based objects.

Images and representation

The focus here is not on the hardware, data standards or mechanics of converting analog data into ones and zeros, although these must be understood as the foundation of all digital representation systems, as discussed by Mostafa (1994). In digital information systems, representation is concerned with the symbol structures used inside a computer. In linguistics, psychology and philosophy, representation involves the broader arena of knowledge representation. In contrast, representation in information science is concerned specifically with document representation. As Lancaster observes, "The main purpose of indexing and abstracting is to construct *representations* of published items in a form suitable for inclusion in some type of *database*" (1991, 1. Emphasis in original.). Smith and Warner (1984) note that the representations used in information retrieval systems may refer not only to documents but also to objects, relationships, and processes. They identify five categories of representation: document representations as surrogates; query representations; document and query relationships; term relationships; and document relationships.

The differences between pictures and representations, as discussed by Wartofsky (1979), provides a definition structure that is useful for considering representation of documents in digital systems. In the physical realm, cultural artifacts are distinguished from naturally occurring phenomena. However, natural objects may be selected as artifacts, as when stones are intentionally sited within a rock garden or shells are collected and arranged for display on a table. Nature may also be approached pictorially, as when a sunset is viewed as a picture. For Wartofsky, all artifacts are representations in that they are symbols that embody intentionality: "a spear is made for hunting and also represents the mode of action of the hunt; a spear picture is not made for hunting but is made expressly as a representation of a spear" (1979, 282). These artifacts may be visual or non-visual. While non-visual artifacts may be auditory or linguistic references, visual artifacts are made to be seen and are not limited to pictures: they may be gestures, hand signs, sculpted objects, kinship resemblances, etc. In this regard, Wartofsky adheres to Goodman's argument that the resemblance requirement of classical representation is not sufficient for defining a representation. Because visual artifacts are culturally constructed and symbolic of intention, it is possible for a picture to not be a representation, but it is equally possible for all pictures to be taken as representations. Following Wartofsky's argument, *picture* will be used in this discussion to refer to visual representations that rely upon or include resemblance; *image* will be used to denote visual representations in general and will not connote reliance on resemblance.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

In the digital environment both pictures and images are made on a plane, and are dependent on line or color. Within the digital environment, the definition of image as a visual representation not dependent upon resemblance is in accord with Twyman's definition of graphic language in that the distinction between verbal and non-verbal elements does not apply. To generate the image on the display monitor, both verbal and non-verbal elements are indiscriminately represented as ones and zeros (the code) and identification of displayed elements as text or image is not a representational issue. Differentiation between text and image occurs only at the level of internal processing respective to the particular digital encoding-decoding system (i.e., relevant to text codes such as ASCII or to image codes such as TIFF or GIF).

Because the present discussion is not about internal systems, whether human or digital, but about external representations, the term image must be understood to subsume all digital displays of text, pictures or graphics. Within this framework, then, the term *picture* will be understood to refer specifically to those displays that convey visual resemblance and contain no text. In the digital domain, *graphic* may refer to any of the following: to text entities as pictures of the alphabet; to any non-text entity; to raster-based, but not vector-based, non-text entities; or to an informational chart but not to a rendered scene. It is used here to refer to displays that contain combinations of text and pictures, or, more typically, to those images which are less dependent on resemblance to the referent, moving toward the purely symbolic. These definitions are not necessarily mutually exclusive; for example, many corporate logos, such as those of IBM and 3M, include alphanumeric characters but are commonly accepted as pictorial conventions. In such cases, their identification as text, picture or graphic is dependent on the application.

Howard (1980) explores the concept of representation by addressing the array of meanings that adhere to the term *represents*. He argues that these various meanings point to the different functions that symbols perform (depiction, description, expression, etc.) and that a particular form of representation, such as language, may actually perform several such functions. Furthermore, each such function may be accomplished through multiple forms of representation. Based on his explication of the terminology and the observation that representation, symbol, and language are commonly equated, he argues that the present dichotomy between visual and verbal representations is both untenable and unrealistic. It is more reasonable, he suggests, to view representation as functional (as in verbs) or symbolic (as in nouns).

Unfortunately, the traditional dichotomy between verbal and visual is reinforced by the fact that symbolic representations are frequently confused with the medium of representation, as in a linguistic description versus a pictorial depiction of an object. In ***Languages of Art*** (Goodman 1976), "the single most elegant and rigorous theory of symbols" (Howard, 1980, 505), Goodman not only contends that resemblance is neither sufficient nor necessary for representation and argues that different ways of presenting information on a surface can imply different ways of relating to those marks. What distinguishes different forms of representation is not so much the visible marks on a surface as the different sets of rules that apply to the re-construction and interpretation of those marks. He contends that the lexical and the functional, the syntactic and semantic, operate in concert to carry specific meaning.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

According to Howard (1980), the cognitive perspective on representation is concerned with the role of symbols in the growth of knowledge and perception in the individual, in communication, and in culture -- in knowledge representation. The cognitive perspective addresses how symbols of different kinds serve to mediate thought and perception. It seeks to identify the psychological, cultural and/or communicative processes involved in the acquisition and use of symbols in thought and action; and it attempts to do so through procedural, propositional, or critical judgment analysis. Given that the computer is a rule-based technology, the focus here is on critical judgment through differentiation -- based upon rules and cues -- and not on the cognitive aspects of knowledge representation but on the rules and cues that result from logical, syntactic and semantic analysis of the representation as symbol.

In the realm of the visual image, Gestalt theory and the Bauhaus movement can provide direction for identification of these rules and cues. While the image as a medium of expression is appropriately explored in terms of sense stimuli as manifested in the field of psychology, it can similarly be explored through application of Gestalt concepts relevant to graphic design.

Gestalt theory

Gestalt is a German word without a direct English counterpart. Loosely identified with form and shape or, more broadly, with essence, Gestalt reflects the concept that "'wholes' are experienced as such and not as the sum of their parts (Krampen 1994, 290). This concept entered discourse in philosophy, psychology, and visual design early in the twentieth century. In psychology, Gestalt theory was applied to perception and sensory organization to explain how the individual sees the world and to identify the relationships that exist between physical patterns in the world and human physiological activity. Kohler (1947) adopted Gestalt theory to counter behaviorist theorists who rejected the study of direct experience, arguing that "when discrete entities unite in a group, the part which equality (or similarity) play in the unification cannot be explained in terms of learning" (84). The application of Gestalt was carried to the extreme in the study of optical illusions by the Berlin school, which held that there was a "direct and lawful connection between physical stimuli and their sensory perception" (Krampen 1994, 291). Cataldo (1966) points out that, over time, differences in the application of Gestalt to perceptual psychology were resolved by focusing on the principle of an isomorphic correspondence between the structural characteristics of visual form and the observed phenomena in human behavior.

Gestalt theory advocates an organismic approach to perception that was in sharp contrast to the more mechanistic approach of traditional science which endeavors to solve problems by isolating elements and reassembling them. Gestalt holds that "there are wholes the behavior of which is not determined by their individual elements, but where the part processes are themselves determined by the intrinsic nature of the whole" (Cataldo 1966, 109). The outcomes of these processes revolve around this "intrinsic nature of the whole" and have been delineated in their subtle variations as a set of Gestalt principles. Of the fourteen principles articulated by Krampen (1994), five are of particular interest here:

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

1. A field naturally tends to become organized into forms.
2. A form tends to be a figure set upon a ground which, in the visual sense, is synonymous with field.
3. Forms vary by their articulation ranging from simple to complex.
4. Similarity and adjacency of forms results in their integration or grouping.
5. The stronger a form, the more meaningful and objective it appears.

In short, Gestalt is theory of fields which posits that strong forms are well articulated, resist disintegration into parts or fusion with other forms, tend toward closure if open, and depend more on the dynamics of central forces than on stimulus properties.

The figure-ground relationship was considered the starting place for perceptual activity (Kennedy 1985) and included principles of good continuation, closure, proximity and similarity. Good continuation refers to the linearity of form that allows the individual to follow a line of text across a page. Closure is related to continuation in that it implies a tendency to close areas or gaps in what is seen as incomplete. Just as it is psychologically more rewarding to achieve closure in activities, closed figures are viewed as more stable.. Proximity, or adjacency, leads to visual grouping of forms based on the nearness of their parts, while the effect of similarity provides that visual parts which resemble each other will tend to be seen as a group. Similarity can involve shape, size, color, direction or even a time period. Thus similarity in time period accommodates the dynamic of film, which incorporates similarity of both time and place. In the Bauhaus in the early decades of the twentieth century, Gestalt principles were applied in education for graphic design and visual communication and formed the basis of stimulus design. Later, Arnheim (1974) extended this work to explore the dynamics of forces and tensions, the central pull-and-push of image design.

As Kennedy (1985) points out, critics of Gestalt have argued that figure-ground is just a pictorial effect -- nothing more than lines and contours on flat surfaces creating edges but not depth. Taking an extremist position, Kennedy argues that, "In principle, perception is simply a means of grasping the underlying mathematics of forms" (1985, 35). Currently, fractal geometry is used to simulate varieties of landscapes as well as other animate and inanimate forms. These simulations are perceived as coastlines, clouds, hills or deserts, yet they are based on geometry and on variations of lines as edges. Another argument against Gestalt derives from the theory that individuals hold in memory schema for types of objects such as noses, houses or even scene arrangements (Gombrich 1960). The extremist Gestalt response is that individuals privilege basic forms and perceive by constructing objects out of the visual building blocks of circle, square, and triangle (Arnheim 1974; Dondis 1973). From this perspective, Gestalt is grounded in perceptual simplicity and relies upon the principles of basic geometry for visual analysis. Indeed, the extreme position that there are basic privileged forms is itself reminiscent of schemata theory.

If these extreme views, which were manifested in the Bauhaus movement, are rejected, then the task of perceiving shape would be unconstrained and therefore boundless given that there are infinite variations in nature. This is not to say that recognition is based first or foremost on perceptual form, as related here to basic geometry. Rather, gestalt theory contributes to the definition of image the principles of perception, continuation, similarity, proximity, and closure. These principles form a geometric basis for visually analyzing the physical world and inform the

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

interpretation of the viewer. Because coding in the digital environment is itself grounded in mathematics, it should therefore be possible to develop coding systems that represent the basic geometric forms of perception.

Gestalts as holistic perceptual structures have also been equated with the concept of sign (see, e.g., Nöth 1990; Krampen 1994). Moles (1966) presents a theory of supersigns or holistic elements of information processing which are themselves made up of subsigns, or sign element building blocks. Nöth explicates Moles's position:

> According to Moles, the perception of a visual image is a process of integrating such subsigns and supersigns within the pictorial whole. More specifically, he proposes a hierarchy of perceptual levels extending from a differential optical element, a geometrical morpheme, a partial image of a signifying object to an iconic phrase and discourse. (Nöth 1990, 451)

Gestalt is thus a system of subsigns within the overall sign system which, as Krampen (1994) observes, provides connections between sign systems. In particular, the "geometric morpheme" is the focus of Gestalt as applied in the Bauhaus movement to the encoding process of image design; but definitive interpretation can never be determined because of the iterative semiotic process of interpretant-representamen-object.

With respect to any given image involved in a semiotic process of interpretation, "the particular properties encoded for a given exemplar are assumed to depend upon the individuals prior knowledge and choice of encoding strategies, the context in which the exemplar is presented, and so forth" (Hayes-Roth, 1977, 323). While semiosis may be the foundation of understanding in specific situations, the goal of an information retrieval system is to address representation at the more generalized level of prototypes. Inevitably, each specific situation or exemplar will vary from this generalization, but constraints must be constructed for efficient technological functioning. Keil has observed that, with respect to cognitive categorization,

> Many of the traditional theories and much of the empirical research on conceptual development have been shown to argue for development away from a holistic mode of representing concepts toward a more analytic mode. The shift is revealed by a decreasing attention over time to all the attributes that typically co-occur with a category and an increasing focus on only a certain subset of attributes that are essential to the meaning or intension of a concept. (Keil 1989, 14)

This shift is acknowledged in the Gestalt idea of the whole as greater than the sum of its parts. Each part can be examined, but the whole is more than any one part or combination of parts. As Chen et al. (1998) have observed, the terms assigned by an indexer are generally more precise and more semantically rich when compared to the product of automatic indexing, perhaps because the indexer understands and interprets the document as a whole rather than focusing on individual words or phrases.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Words and pictures

In art history, the tendency has been to emphasize the rhetoric of iconography (i.e., meaning) over formal analysis (i.e., structure) because the latter is strongly associated with aesthetics (Vaughn, 1992). Similarly, in discussions of image retrieval in information systems, the focus has been on the use of verbal content and concept-based representations: the *of*-ness and *about*-ness of the image (Rasmussen, 1997; Layne, 1994; Jörgensen, 1996; Small, 1991). There are, however, inherent problems with the verbal representation of images, problems that are exemplified by consideration of the painting *Las Meninas* by Velasquez. Foucault (1970) suggests that this work exemplifies "the representation of classical representation" (16) in that painting, up to the sixteenth century, attempted to imitate space. Searle (1980) elaborates on the problems associated with the "excruciating" realism of the surface features of *Las Meninas*, pointing out that "the firm ground of pictorial realism begins to slip away" (250) when the viewer considers the reflection in the mirror on the back wall of the painting. Thus, for the causal viewer, this painting seems to be of the artist painting a canvas and of the young girl, her attendants and dog. But the canvas in the painting is not visible to the viewer and the mirror's reflection of the King and Queen of Spain suggests that the painting is actually about the royal couple. Searle suggests that the surface features are indeed the classical representation but that these visual surface features introduce another level of representation that combines resemblance and point of view with the more traditional visual aspects. He concludes that "there is no way to answer the question What is the picture a picture of? that does not include reference to the picture" (258).

With respect to representation of images in electronic retrieval systems, there has been little discussion of the relationship between the physicality of the image and its semantic attributes. As Eco (1996) observes, "even if it were true today that visual communication overwhelms written communication, the problem is not to oppose written to visual communication; the problem is how to improve both" (298).

Representation in information retrieval systems

There is, in fact, no body of literature that specifically addresses document representation. Rather, there is a large body of literature dealing with classification or knowledge organization which covers issues of representation, including classification schemes, thesaurus construction, indexing and abstracting. The terminology of this literature is often confusing, ambiguous, and imprecise: the term "classification", for example, is used to refer broadly to any process for organizing knowledge or, more narrowly, to the specific order of well-defined and mutually exclusive classes (Jacob 1992). For the purposes of this discussion, the term "document representation" will be used to refer to the wide range of indexing techniques which includes faceted and enumerative classification schemes, subject heading lists, indexing by assignment from controlled vocabularies, and indexing by extraction of keywords, as well as full-text indexing.

The components of traditional document representation include bibliographic data, assigned descriptors or keywords, location codes, abstracts and, more recently, full-text. For non-text objects such as pictures and video, the retrieval system generally relies on text-based representations. Furthermore, there is often no clear distinction between the data and the

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

document or object for which it serves as surrogate. In the present context, *data* will refer to a record that is stored in a computer, whether that data points to a single document or is grouped with other data to form an aggregate representation of a collection of documents. Furthermore, it is assumed that all documents are ultimately objects -- or things -- in the real world. These objects might be books, which are textual even if they contain pictures, or they might be non-text objects such as videos, physical objects, images, or anything else that has the potential to be informative. The definition of objects as things in the real world with the potential to be informative would account for the digital image which, although an object made up of an aggregate of on-off bits, has no referent in the real world but does have the potential to be informative.

Practical and theoretical research dealing with document representation of non-text materials tends to concentrate on text-based approaches, ranging from development of the *Art and Architecture Thesaurus* to discussions of appropriate levels of subject description (e.g., Krause, 1988; Layne, 1994). For example, an oil on canvas by Raphael (physical representation) might be a picture *of* a woman and an angel (content representation) that is *about* the Annunciation (concept representation). If non-textual information is to be represented only through text, then current text-based approaches to information systems will continue to focus on text retrieval methodologies. Research on pattern matching points to the feasibility of an image-specific approach to information systems design. Through pattern matching, the raster portion of one image can be matched against that of other images or patterns of texture can be identified and compared (Caivano, 1994). Unfortunately, however, the notion that pattern matching "is really the visual equivalent of the 'word search' that is a standard feature of every word-processing and database package" (Vaughn 1992, 2) has led to exploration and application of traditional text-based algorithmic approaches (Gouge, 1996; Hibler, 1992) without consideration of how non-textual data is to be represented in the system.

Visual information systems

Over the last decade, there have been a number of proprietary or collection-centered attempts to develop visual information retrieval systems, many of which originated with art and museum collections. Morelli is one such system that relies on analysis of form, as suggested by Small (1991). Morelli matches, sorts and orders pictures on the basis of visual characteristics which have been derived from the digitization process (Vaughn, 1992). Stenvert (1992) describes Morelli as an initial attempt to promote a scholarly, morphological study of images using a calculated abstraction based upon the bitmap:

> It must be remembered that the digital image does in itself perform an analysis on an image. It divides it into a series of units. As it happens, the categories used for these do correspond to the three basic areas of pictorial experience mentioned by Renaissance art theorists, namely design, color, and tonality (or chiaroscuro). This approach results in a means with which, for the first time, art historians would have ways of classifying pictures visually, rather than verbally. (Stenvert 1992, 28)

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

The Morelli system is concerned with providing an objective means for describing and identifying pictorial characteristics, such as form, configuration, motif, tonality and color. It carries out an image search by matching and comparing pixels to identify pictorial characteristics. This is carried out at ever more granular matrices of pixels, down to the level of 4x4 pixels. Each matrix records a value for configuration (edge clarity), for tonality (light), and, in the near future, for color. Form identification is thus based at the pixel level of object shapes within the image rather than from relationships among objects, as suggested by the notion of physical arrangement.

Wide-spread access to the World Wide Web, however, in concert with the recent rapid advances in imaging technology, has focused attention on more generalizable systems such as QBIC, Virage, and VisualSEEk. The goal of these systems is simple: to provide a manageable number of images for browsing. As described by Gupta, Santini and Jain (1997), retrieval in such systems is based on visual object appearance involving four categories of information: features, feature space, feature groups and image space. These categories are defined within the domain of online system design. Thus a *feature* is "a derived attribute obtained by transforming the original visual object through an image analysis algorithm; it characterizes a specific property of an image" (37). For example, a distance algorithm computes a difference value between features or feature groups: the greater the distance, the less the match. *Feature space* is the region inhabited by a specific image property, or appearance feature, such as texture, color or a bounded object. The feature space is treated as an information object which can be queried by a number of operations: find boundary, select by spatial constraint, select by distance, find nearest neighbor *k*, etc. In addition, simple features can be grouped into more complex features and operated on by distance aggregation which allows feature joining (e.g., color and texture in red fabric) or by dimension joining which provides for creation of a set of feature primitives (e.g., a red-white stripe detector for the American flag). It is possible that these techniques may provide the algorithmic building blocks for an image retrieval system that will operate at a more complex level of physical arrangement.

Image space, on the other hand, considers the image as a totality and attempts to extract its regions based on appearance features. This approach provides for computational comparisons once a feature has been represented numerically; but there is, as yet, little consensus as to what features to derive in the first place. Color is an obvious first one to try, as demonstrated by Virage and QBIC: if the user is looking for a picture of a garden, the query might be for images that show 20% blue on top, 40% green on bottom, and a pink region in the middle (Gupta, Santini, & Jain, 1997). Both QBIC and Virage currently allow for global feature queries, but they do not support queries based on distance algorithms or spatial relationships. Jain (1997) describes these systems as examples of "query by pictorial example". But when the image database is large, comparative computations become prohibitive and indexing becomes essential. QBIC uses image-only attributes for indexing: for example, shape is defined only for individual segments and is not satisfactory for measures on complete images. Other attributes are color and texture; but neither of these relies on domain-level knowledge, which Twyman (1979) argues may be an important factor in both recognition and organization of images. He suggests the need to identify those attributes which are important for domain objects and to determine if they are better represented using image alphabets or image objects.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Shape research

After considering existing content-based visual query systems such as Virage, QBIC and VisualSEEk, Chang et al. (1997) developed WebSEEk, "a semiautomatic image search and cataloguing engine" (66) in which the image can be represented at several different levels: at the feature level (e.g., color, texture, shape); at the object level (e.g., foreground object); at the syntax level (e.g., video shot); and at the semantic level (e.g., subject of the image) Obviously, image retrieval systems differ in their effort to categorize visual information into semantic ontologies: thus, while current image search engines such as Virage and QBIC support querying by features, they rely on different feature sets and methods. Content-based approaches are not intended to replace the keyword approach but are viewed as complimentary. Such an approach generally starts with a rich set of low-level, generic visual features and attempts to derive a higher-level semantics by applying domain knowledge provided by users or applications. VisualSEEk allows users to search by localized color regions and spatial layout: sets of colors and distances between colors are then indexed as "near-symbolic images which lend themselves to efficient spatial query" (Chang et al., 1997, 89). To enter a query, the user sketches regions and assigns boundaries and properties.

But are there conventions of representation or boundaries that could be employed to serve the same purpose? Twyman suggests that there are image feature schemas -- existing conventions of representation and boundaries -- that could facilitate the standardization of image retrieval systems. The existence of feature schemas has been explored in algorithmic shape research (Forsythe et al.. 1997; Corridoniet al. 1998) and through the development of taxonomies (Lohse et al. 1991, 1994; Twyman, 1979).

Forsyth et al. (1997) describe promising research for finding pictures of objects based on spatial attributes. They argue that manual, text-based indexing is unrealistic for very large image databases and propose a model that uses grouping rules based on object domain criteria. In their experiments, an object is modeled as a loosely coordinated collection of grouping and detection rules that reflect "both surface properties (pattern) and shape information" (136). This approach works with objects whose boundaries are potentially variable since object identification is based on non-geometric cues such as color and the relationships between components of the image. Although this approach would probably not be appropriate at the level of specific individuals, it should be effective at more general levels of identification and organization. The key here is the identification of objects and their domain specific constraints. Thus, for example, to find images containing naked people, the model "first locates images containing large areas of skin-colored regions, then within these areas, finds elongated regions and group them into possible human limbs and connected groups of limbs" (Forsyth et al., 1997, 132) The object in this example is the human figure and its constraints are skin-color and the spatial grouping or configuration of arms, legs, torso, etc. Grouping rules are built from local (skin color areas) to global (size and shape of areas), and from generic (limb shape) to specific (limb and torso articulations).

Forsyth et al. critique current object recognition systems for not coding "spatial organization in a way that supports object queries"(1997, 122). Current approaches are frequently based either on modeling precise geometric forms or on creating hierarchies of rectangles of internal color on a grid (e.g., QBIC and Virage). In the first situation, if the side view of a horse is modeled

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

geometrically, the front view would be unrecognized; in latter approach, a query for images with a large proportion pixels of a particular color would retrieve a set of images whose content was potentially variable since the connection between content and color must be inferred by the individual generating the query. It may be possible, however, to evaluate images on the basis of Gestalt principles and configurations to determine if there are domain objects that could be used for building grouping rules.

Image taxonomies

Smith and Warner (1984) observe that, "In developing a taxonomy, one must be concerned with both classification and nomenclature, where classification is the process of establishing and defining systematic groupings and nomenclature is the allocation of names to the groups so produced" (123). While some would argue that taxonomies indicate mutually exclusive categories and that categories are more flexible structures, allowing overlap or choice in data assignment (Jacob, 1992), both approaches are tools for making generalizations about the data under consideration.

There is a small group of researchers who have attempted categorizations of various aspects of graphic language documents. Bertin (1983/1967) is concerned with the construction of diagrams, networks and maps, or, as they are known today, info-graphics or quantitative displays. He identifies eight variables for non-moving graphic representational systems: a visible mark expressing a pertinent correspondence can vary in relation to the two dimensions of the plane, size, value, texture, color, orientation and shape. In one example, Bertin offers 100 different variations when these variables are applied to a single data set. In contrast, Caivano (1994) works with the single variable of texture. He describes a system for ordering textures based on other variables such as size, saturation, density and organization which, he argues, can be used to develop systematic and algorithmic descriptions of textures. Tufte (1983, 1990, 1997) also focuses on quantitative graphics. Although he offers guidelines for the production of graphic language documents, these guidelines are intuitively derived from exemplar documents and are therefore very general in nature. He provides no schema or categorization structure.

Lohse (Lohse et al, 1991, 1994) is concerned with how people code and decode data in "informational graphics" -- in visual images used for information communication. His research explores how people classify graphical information, by which he presumably means "graphs", into meaningful, hierarchically-structured categories. In contrast to earlier approaches in which the researcher generated a system of classes or categories (e.g., Twyman, 1979), Lohse (Lohse et al. 1991) asked his subjects to generate categories for a collection of images. In a follow-up study (Lohse et al., 1994), he confirmed the validity of his categorical structure and, using multidimensional scaling, identified specific features that characterized each of the ten categories:

1. Graph -- encodes quantitative information using position and magnitude of geometric objects.
2. Table -- an arrangement of words, numbers, signs, or combinations of these intended to exhibit a set of facts or relationships in a compact format.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

3. Time chart -- display of temporal data.
4. Network chart -- shows relationships among components.
5. Structure diagram -- provides a static description of a physical object.
6. Process diagram -- describes the relationships and processes associated with one or more physical objects.
7. Map -- a symbolic representation of physical geography.
8. Cartogram -- spatial maps that show quantitative data.
9. Icon -- carries a single interpretation or meaning for a picture.
10. Photo-realistic picture -- realistic image of an object or scene.

Although Lohse's work represents a rare experimental approach to the question of graphic taxonomies, he further confuses the research arena by his failure to define his document type and by his inconsistent use of terminology.

Rankin (1989 is also concerned with the categorization of graph types, but he draws a very narrow focus on only those graph forms which use an *xy* coordinate space that is both explicitly stated and numerically scaled. He distinguishes between functional and structural classifications, placing his own work in the latter grouping along with that of Doblin (1980) and Twyman (1979). Functional classifications are based on the intended use of the graphic documents, while structural classifications focus on form irrespective of use. Doblin's work, however, is potentially more complex than that of Rankin or Twyman in hat, once he establishes a grid based on form and content, he then attempts to map the results to the purpose of the message.

Twyman (1979) proposes a matrix for categorizing graphic languages, or visible communications, that is based on practical application (or function) in combination with form. His matrix presents a theoretical approach to grouping in terms of two variables: method of configuration and mode of symbolization. Method of configuration is the graphic organization or physical structure of the image. Mode of symbolization refers to the representational form of the image, which may be verbal/numeric (i.e., numbers or other conventional characters such as typographic letters), pictorial or a combination of verbal/numeric and pictorial. He further divides pictures into schematics, which are "all purposeful graphic marks that are not words, numbers, or pictures" (Twyman, 1985, 247). This, he admits, is a fuzzy division and there can be much overlap between picture and schematic.

Acknowledging that all documents carry a message as described by Doblin (1980), Twyman (1979) observes that the internal graphic organization or structure of the image -- the method of configuration -- "influences and perhaps determines the searching, reading, and looking strategies adopted by the user" (121), thus building a theoretical position similar to that which sees the principles of Gestalt as the foundation for visual organization. Because he can find no accepted terminology for this approach, he creates his own based on the fundamental notion of linearity, but he fails to adequately define linearity and the resulting terminology relies, instead, on his selection of visual exemplars.

With respect to the specific cells, or categories, in his matrix, Twyman suggests that the verbal/numerical linear-interrupted category, best exemplified by traditional texts, is the dominant form precisely because of its easy production. He also suggests that commonly used categories

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

such as this work well by virtue of the simple fact that they are commonly used, implying the notion of convention as suggested by Riegl (1992). This could be changing within the environment of the world wide web, however, as alternative methods of visual representation are more readily available to and easily adaptable by non-specialists. But the question of the relationship between conventions of configuration and domain-level knowledge remains: How does the viewer develop an appropriate strategy for extracting information once he has identified the method of configuration? And can the method of configuration itself be coded in the representation in a digital system?

Twyman attempts to isolate the various encoding options and build an analysis of their application, an approach indicating a structuralist or process-based perspective that would generally be identified with formalism in art theory since it considers only the visual elements and properties of an image. Carroll (1982) points out, however, that, while "No one doubts that there is a structure to visual communication", these structures may be manifested in a multiplicity of forms that can encompass "innate structure, semiotic structure, grammatical structure, cognitive structure, the structure of eye movements, even the structure of convention" (371).

Communication as function

The function of an image can be broadly defined as communication, but the precise nature of this communication depends upon the interests and purposes of the discussants -- of the encoder and the decoder. Communication is variously described as intending to convey knowledge; to inform interpersonal relationships; to persuade; to express beauty and/or emotion, etc. (Barthes 1985; Larsen 1985; Doblin 1980; Twyman 1979, 1985). Designers -- whether they are designers of persuasive advertising or of expressions of beauty -- are thus encoders of visual communications and are intimately concerned with understanding how visual phenomena are used in the act of communication. Adopting the perspective of an encoder, Berger (1989) outlines five interacting factors that affect visual communications: the artist, or encoder, who creates the image; the audience, or decoder, who receives the image; the work of art which is an image itself but may actually consist of a number of elements which are images in their own right; the social context, whether that within which the image is encountered or that within which the image has been created; and the medium.

Structure as Form

Although spatial dynamics -- the geometry of the Gestalt involved in interpreting the physical world -- provide a foundational relationship between individuals and the representations and creations of objects in the world in which they live and move, they are frequently taken for granted. Bragdon (1922) points out that, in the microscopic world of nature, the constructing units tend to arrange themselves with relation to simple geometrical forms -- the visual building blocks of circle, square, and triangle. These forms give rise to unity and simplicity in the face of complexity: the honeycomb, the snowflake, the proportions of the human figure, the triangular arrangement of the features of the human face, etc.. Boulding (1968) suggests a typology of

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

images beginning with the simplest level of organization, called static structures, which are exemplified by the statue, the picture, or, more abstractly, the atom. Successive dimensions of organization are identified as mechanical, homeostatic, biologic, botanical, animal, and human; but, because Boulding's typology is agglomerative, "Each level of organization includes characteristics from all the lower levels" (28).

There are numerous terms that describe the concept of form, or structure, as applied to images. In everyday conversation, many of these terms are used interchangeably; but, within specific domains, individual terms are applied more carefully. In graphic design, the terms *shape* and *form* demonstrate specific differences in that *form* implies the three dimensions of length, width and depth, while *shape* indicates only length and width (Ragans, 1988). According to Hurlburt (1977), the publishing activity of page layout exemplifies the notion of form defining structure. Focillon (1948) contends that, in fine art, form is the "modality of life", a graph of activity that is inscribed in space and time.

Barthes (1985) explores the relationship of form to meaning in a discussion of Eisenstein's film "Ivan the Terrible". The specific frame, or still, that he discusses shows two courtiers showering the young tsar's head with gold. Barthes distinguishes three levels of meaning in this frame: the informational or communication level; the symbolic level; and a third level that he labels the obtuse. On the informational level, Barthes finds meaning in the setting, the costumes, the characters, and their relationships. On the symbolic level, he finds that the shower of gold conveys the theme of wealth and a more general signification of exchange. The obtuse level is focused on the signifier or representamen -- the physical features found in the visual image itself which can only be described by the density of the courtiers' makeup, the delicate line of one courtier's eyelid or the shape of the "stupid" nose on the other. These are the structures of the image -- the forms or shapes -- selected by the filmmaker as encoder. They are the signs that contribute to the Gestalt of the frame. Furthermore, the obtuse meaning is not semantically situated but "remains suspended between image and its description" (Barthes 1985, 55). The implication here is that there can be no linguistic description of this meaning, suspended as it is between the visual and the verbal. However, the principles of Gestalt might be brought to bear to explicate the relationship between communication and structure -- between function and form. Focillon (1948) captures the essence of this approach in his statement regarding the principle of form: "the sign bears general significance, but having attained form, it strives to bear its own individual significance; it creates its own new meaning, it seeks its own new content, and then endows that content with fresh associations by the dislocation of familiar verbal modes" (5).

The iconography of images is described by Focillon (1948) as having three categories: variation of form with the same meaning; variation of meaning with the same form; and form devoid of meaning. He illustrates the third category -- form devoid of meaning -- with the history of the interlace form which has its origins in the symbol of entwined snakes in the caduceus of Aesculapius, the Roman god of medicine and health. He points out that the symbolic association with medicine was lost, however, as the interlace form was incorporated into the ornaments and monumental architecture of Islam and East Christian communities and took on a life of its own. In information retrieval systems, however, the iconography of images is limited to the first two of Focillon's categories and involves the naming of entities and/or their component parts. This

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

process serves as the basis for representational languages in most information retrieval systems that include images.

Although physical arrangement in space is ubiquitous in all human relationships and occurs as an attribute of creations and/or interpretations across all dimensions of reality, physical arrangement of entities within an image document is a spatial attribute that is frequently neglected in systems of representation. Broadly speaking, it is easy to look at pictures and assume that they can be represented in the same fashion as alphanumeric documents. Reliance on existing schemes devised for alphanumeric documents, however, overlooks the graphic nature of pictures and their inherent spatial physicality. Although it has been argued that no one scheme can order reality for everyone (Jacob & Albrechtsen, 1997, 1998), current attempts at the development of standards for representing collections of pictures are driven by the verbally-oriented ethnocentrism of the current classificatory workforce. Such a representational system closes the door to future interpretations that might actually be based upon physical arrangement given that current methods of physical description generally fail to incorporate representations of a picture's internal spatial attributes or geometric structure. If objects are represented solely by linguistic interpretations of conceptual meaning generated within such a verbally-biased context, then objects residing in both public and private collections are at risk of becoming unretrievable.

Form and function: The Bauhaus

Focillon (1948) argues that "A work of art is an attempt to express something that is unique, it is an affirmation of something that is whole ... but it is likewise an integral part of a system of highly complex relationships" (1). If the phrase "work of art" is replaced with "image", as defined above, then Focillon's statement reflects the immediate challenge confronting the design of digital image retrieval systems. This challenge involves issues of representing -- of naming or categorizing for subject access -- non-verbal materials (Jörgensen 1996; Layne 1994; Small 1991). All images are, at a minimum, a slice from the real world, from the selection of a plot of ground to produce an aerial photograph to the selection of what to display within the confines of the medium's viewport, be it at the book page, television screen or computer display. All pictures are at once complex but nonetheless simple in their wholeness: that is, all images integrate the complexity of that which has been selected within the unique perspective of the selector as encoder.

In the twentieth century, one major approach to this process of selection had its roots in the Bauhaus, a progressive school of art founded in Germany by Walter Gropius in 1919. Cataldo (1966) has observed that the Bauhaus ushered in a new era in graphic, industrial, and architectural design: it emphasized social (or self-) consciousness; valued craftsmanship at all stages of production; and not only crystallized but systematized the industrial age, accepting the machine as the essential vehicle of form by combining quality craftsmanship with art. Lupton and Miller (1993) acknowledge the mythic proportions of the Bauhaus as the foundation of modern design. As such, the Bauhaus advocated the search for a unified work of art centered around a system of signs that were both natural and universal -- a language of vision and a fundamental visual grammar whose central elements were the triangle, the circle, and the square. But Hurlburt (1977) contends that "No movement of modern design has had more written about it - and probably no

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

movement has been so misunderstood than the Bauhaus" (38). Indeed, Wolfe (1981) humorously attacks its influence in architecture by asking "has there ever been a place on earth where so many people of wealth and power have paid for and put up with so much architecture they detested?" (7).

Misunderstanding of the Bauhaus can be traced to several sources. In the first place, the language of vision was understood as speaking directly to the eye, following Gestalt theory as it was manifested in the physiological optics of the Berlin school. Secondly, although the term "graphic" was acknowledged as relating to both writing and drawing as well as to conventions of data display in the sciences, its relationship with verbal language was only analogical. Nonetheless, subsequent research has focused on the ability of visual languages to replicate the forms and functions of verbal languages (Carroll 1982; Knowlton 1966; see also Kolers et al. 1980). These efforts continue in the arena of digital systems with the work of Cobb and Petry (1998), Chang et al. (1986), Caivano (1994) and others. For the most part, current research attempts to devise algorithmic grammars for the geometry of image processing: existing digital systems currently match alphanumeric strings and edge matching in images would appear to be the logical next step.

The popular legacy of the Bauhaus is contained in the clichés that purportedly originated with two of its proponents: Sullivan's "form follows function" and van de Rothe's "less is more." Sullivan was an American architect who advocated the design principle that "a building should publish to the beholder the purpose which it subserves -- what it is for, what it is about, why it is as it is and not otherwise" (Bragdon, 1932, 5). Bragdon's accounts of his interaction with Sullivan demonstrate that, by "function", Sullivan meant the natural characteristics of the materials themselves and not the sociocultural application or utility of the final product. Thus, for example, a building using steel-framed construction materials of the industrial era should not be given the appearance of a solid masonry structure. Rather, the form should follow the function of the materials themselves and the techniques employed in construction using those materials.

Sullivan's legacy -- the notion that form follows function -- has been removed from its original context and its intended meaning has consequently been lost. The notion that function should determine form was intended to address the need for architecture to take into account the unique characteristics of the construction materials themselves. Instead, popularization of "form follows function" led to misinterpretation of Sullivan's intent and subsequent neglect of the material characteristics of the medium in favor of the utility of the end-product. This misinterpretation of the notion of function can be extended to the representation of images in digital retrieval environments in that the verbal description of the image is privileged over the graphic form -- the visual structure or geometry of the image.

The Bauhaus suggests that form follows function -- that the concepts of form and function are closely intertwined in the creative design process. Function must be understood to refer to both the capacities of the materials to be used in the creative process and the intended use of the created object. The nature of the material (form) of the product impacts design of the functional form (use); but the intended use drives the selection of functional form (material). The product is the result of the interplay between the material form and the functional form. Within this

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

framework, then, form implies the union of theory, materials and techniques available for manifestation of the object/product in a particular style by the particular artist. But can these principles of the Bauhaus be usefully applied to information retrieval in digital image collections? And what does that mean?

In the art world, the Bauhaus ushered in abstract expressionism and constructivism by considering the underlying structure of the image and the idea as being created in materials. It is the function of the materials that determines the form of the idea; and the form -- the theory, materials, and techniques -- is the embodiment of the idea as expressed through the medium of the material. The craft is the manifestation of form within the style of the creator. Can the embodiment -- the unity -- be captured as a retrieval tool through such techniques as shape algebras? How is shape algebra -- as a material (digital) and technique (algebra) -- to be translated into a verbal expression of the idea? It is necessary to identify ways of linking shape algebra, which explores the unity of the image, to verbal translations of the process of idea embodiment. According to Arnason (1968), Veltrusky suggests that this translation process is illustrated in the semiotic effects of the variable of color (e.g., warm, cool, etc.). But if, as Mukarovsky (1976) observes, the form of the idea is not attended to, then "the study of the its structure will remain necessarily incomplete" (quoted in Arnason, 1968, 8), doomed to be interpreted as either purely formal structure or psycho-physiological states.

Conclusion

The Bauhaus suggests there exists a unity of theory, material and idea that cannot be captured by merely verbalizing descriptors that would represent each of these aspects. Jörgensen (1996) and Small (1991) both suggest that the variable of location is inadequately represented verbally. Approaches within information retrieval currently focus on function as use -- on the subject or content; but such representations are necessarily incomplete. Gestalt suggests a way to look at form through shape algebras which could be linked to ideas and content within bounded domains to create a more comprehensive representation tool.

The migration of classification theory and practice into the digital environment opens new possibilities for representation. The current discussion would not be useful if it were not for growing interest in the development of tools that can be applied to the retrieval of digital document representations. Although the mathematically-based communication model of Shannon and Weaver is integral to the mechanics of digital activities, it has been rightly criticized for inhibiting the development of information processing systems by focusing too rigidly on a single modality (Sonesson, 1995). Because the essential characteristic of the digital medium is mathematical, application of Gestalt principles in line with an enlightened understanding of how "form follows function" points to the potential for retrieval informed by the internal structural characteristics of the image -- retrieval techniques informed by structural characteristics that would utilize the potential for geometrical recognition. Just as structure and function are intertwined, however, so are geometry and representation: each contributes an important thread to the texture of description.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

References

Arnason, H. H. (1968). *History of modern art*. Englewood Cliffs, NJ: Prentice Hall.

Arnheim, R. (1974). *Art and visual perception- a psychology of the creative eye- the new version*. Berkeley: University of California Press. (first edition 1954)

Barthes, R. (1985). *The responsibility of forms, critical essays on music, art and representation*. (R. Howard, Trans.). New York: Hill and Wang.

Berger, A. A. (1989). *Seeing is believing*. Mountain View, CA: Mayfield Publishing.

Bertin, J. (1983). *Semiology of graphics*. (W.J. Berg, Trans.). Madison: University of Wisconsin Press. (Original work published 1967.)

Boulding, K. (1968). *The image, knowledge in life and society*. Ann Arbor: University of Michigan Press.

Bragdon, C. (1922). *The beautiful necessity*. New York: Alfred A Knopf.

Bragdon, C. (1932). *The frozen fountain*. New York,: Alfred A. Knopf.

Caivano, J L. (1994). Towards an order system for visual texture. *Languages of Design, 2*, 59-84.

Carroll, J.M. (1982). Structure in visual communication. *Semiotica, 40*(3/4), 371-392.

Cataldo, J. W. (1966). Graphic design & visual communication. Scranton, PA: International Textbook.

Chang, S-K., Ichikawa, T., & Ligomenides, P.A. (Eds.). (1986). *Visual languages*. New York: Plenum.

Chang, S-F, Smith, J. R., Beigi, M., & Benitz, A. (1997). Visual information retrieval from large distributed online repositories. *Communications of the ACM, 40*(12), 63-71.

Chen, H., Martinez, J., Kirchhoff, A., Ng, T. D., Schatz, B. R. (1998). Alleviating search uncertainty through concept associations: Automatic indexing, co-occurrence, analysis, and parallel computing. *Journal of the American Society for Information Science, 49*(3), 206-216.

Cobb, M. A. & Petry, F. (1998). Modeling spatial relationships within a fuzzy framework. *Journal of the American Society for Information Science 49*(3), 253-266.

Corridoni, J.M., Del Bimbo, A., & Vicario, E. (1998). Image retrieval by color semantics with incomplete knowledge. *Journal of the American Society for Information Science, 49*(3), 267-282.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Doblin, J. (1980). A structure for nontextual communication. In P. Kolers, M. Wrolstad, & H. Bouma (Eds.), *Processing visible language 2* (pp. 89-111). New York: Plenum.

Dondis, D. A. (1973). *A primer of visual literacy*. Cambridge: MIT Press.

Focillon, H. (1948). *The life of forms in art*. 2nd rev. ed. (C. B Hogan & G. Kubler, Trans.). New York: George Wittenborn, Inc. (Original translation published 1942)

Forsyth, D. A., Malik, J., Leung, T. K., Greenspan, H., & Fleck, M. M. (1997). Finding pictures of objects in large collections of images. In P.B. Heidorn & B. Sandore (Eds.), *Digital image access & retrieval* (pp118-139). Urbana-Champaign: Graduate School of Library and Information Science at the University of Illinois.

Foucault, M. (1970). *The order of things*. New York: Vintage Books.

Gombrich, E. H. (1956). *Art and illusion.* Princeton, NJ: Princeton University Press.

Goodman, N. (1976). *Languages of art: an approach to a theory of symbols*. Indianapolis: Hackett.

Gouge, J.O. & Gouge, S.B. (1996). Is the pixel-based model for image processing simply wrong? *Advanced Imaging*, 10-13, 66.

Gupta, A., Santini, S., & Jain, R. (1997). In search of visual information in visual media. *Communications of the ACM, 40*(12), 35-42.

Hayes-Roth, B., & Hayes-Roth, F. (1977). Concept learning and the recognition and classification of exemplars. *Journal of Verbal Learning and Verbal behavior, 16*, 321-338.

Hibler, D. J.N., et al. (1992). A system for content based storage and retrieval in an image database system. In *Image Storage and retrieval systems: Proceedings of the International Society for optical Engineering* (pp. 80-92). San Jose, CA.

Howard, V.A. (1980). Theory of representation, three questions. In P. Kolers, M. Wrolstad, & H. Bouma (Eds), *Processing visible language 2*, (pp. 501-516). New York: Plenum.

Hurlburt, A. (1977). *Layout, the design of the printed page*. New York: Watson-Guptill.

Jacob, E. K.. (1992). Classification and categorization : drawing the line. In: B. Kwasnik, & R. Fidel, R. (Eds.), *Advances in Classification Research: Volume 2*. Medford, NJ: Learned Information for the American Society for Information Science.

Jacob, E. K., & Albrechtsen, H. (1997). Constructing reality: the role of dialogue in the development of classificatory structures. In *Proceedings of the 6th International Study Conference on Classification Research.*

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Jacob, E. K., & Albrechtsen, H. (1998). When essence becomes function: poststructuralist implications for an ecological theory of organizational classification systems. Information Seeking in Context: Proceedings of the 2nd International Conference on Information Needs, Seeking and Use in Different Contexts, Sheffield, England, August 1998.

Jain, R. (1997). Visual Information Retrieval in Digital Libraries. In P.B. Heidorn and B. Sandore (Eds.), *Digital image access & retrieval,* (pp. 68-85). Urbana-Champaign: Graduate School of Library and Information Science, University of Illinois.

Jörgensen, C. (1996). The applicability of selected classification systems to image attributes. Advances in Knowledge Organization, 5, 189-197.

Keil, F. C. (1989). *Concepts, kinds, and cognitive development.* Cambridge, Ma: The MIT Press.

Kennedy, J. M. (1985). Arnheim, Gestalt theory and pictures. *Visual Arts Research 11*(1), 23-44.

Knowlton, J. Q. (1966). On the definition of picture. *AV Communication Review, 14* (2), 157-183.

Kohler, W. (1947). *Gestalt psychology.* New York: New American Library.

Kolers, P., Wrolstad, M. & Bouma, H. (Eds). (1980). *Processing visible language 2.* New York: Plenum.

Krampen, M. (1994). Gestalt. In T.A. Sebeok, (Ed.), *The encyclopedic dictionary of semiotics, tome 1* (pp. 292-293). Berlin: Mouton de Gruyer.

Krause, M. G. (1988). Intellectual problems of indexing picture collections. *Audiovisual Librarian, 14*(2). 73-81.

Lancaster, F. W. (1991). *Indexing and Abstracting in Theory and Practice.* Champaign, IL.: University of Illinois Press.

Larsen, R.E. (1985). What communication theories can teach the designer of computer-based training. *Educational Technology,* (July)16-19.

Layne, S. Shatford. (1994). Some issues in the indexing of images. *Journal of the American Society for Information Science, 45*(8),583-588.]

Lohse, G., Biolsi, K, Walker,N., Reuter, H. (1994). A classification of visual representations. *Communications of the ACM, 37*(12), 36-49.

Lohse, G., Walker,N., Biolsi, K, & Reuter, H. (1991). Classifying graphical information. *Behaviour & Information Technology, 10*(5), 419-436.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Lupton, E. & Miller, J.A. (Eds.). (1993). *The abcs of [triangle, square, circle]: the Bauhaus and design theory.* New York: The Cooper Union and Princeton Architectural Press.

Moles, A. (1966). *Information Theory and Esthetic Perception* (J. E. Cohen, Trans.). Urbana: University of Illinois Press. (Original work published 1958).

Mostafa, Javed. (1994). Digital image representation and access. In M.E. Williams, (Ed.), *Annual Review of Information Science and Technology (ARIST). 29*, 91-135.

Mukarovsky, J. (1976). Art as semiotic fact. In L. Matejka & I.R. Titunik (Eds.), *Semiotics of art, Prague school contributions* (pp.3-9). Cambridge, MA: MIT Press.

Nöth, W. (1990). *Handbook of semiotics.* Bloomington: Indiana University Press.

Ragans, R. (1988). *Arttalk.* New York: Macmillan/McGraw-Hill.

Rankin, R. O. (1989). The development of an educational design model. *Education Technology Research and Development 37*, 25-46.

Rasmussen, E. M. (1997). Indexing images. In M.E. Williams, (Ed.), *Annual review of information science and technology (ARIST), vol. 32*, (pp. 169-196).

Riegl, A. (1992). *Problems of style, foundations for a history of ornament.* (E. Kain, Trans.) Princeton, N.J.: Princeton University Press. (Original work published c.1893)

Searle, J.R. (1980). Las Meninas and the paradoxes of pictorial representation. In W.T.J. Mitchell (Ed.), *The language of images* (pp. 247-258). Chicago: University of Chicago Press.

Small, J. P. (1991). Retrieving images verbally: no more key words and other heresies. *Library Hi Tech 33*(1), 51-60+67.

Smith, L. C. & Warner, A. J. (1984). A taxonomy of representations in information retrieval system design. *Journal of Information Science, 8*(3), 113-121.

Sonesson, G. (1989). *Pictorial concepts.* Lund, Sweden: Lund University Press.

Stenvert, R. (1992). Bridging the gap between pixel and picture. *Computers and the history of art, 2(*2),19-24.

Tufte, E. R. (1983). *Visual display of quantitative information.* Cheshire, CT.: Graphics Press.

Tufte, E. R. (1990). *Envisioning Information.* Cheshire, CT.: Graphics Press.

Tufte, E. R. (1997). *Visual explanations, images and quantities, evidence and narrative.* Cheshire, CT.: Graphics Press.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

Twyman, M. (1979). A schema for the study of graphic language (tutorial paper). In P.A. Kolers & M.E. Wrolstad, (Eds.), *Processing of visible language volume 1*, (pp. 117-150). New York: Plenum Press.

Twyman, M. (1985). Using pictorial language: A discussion of the dimensions of the problem. In T. Duffy & R. Waller, (Eds.), *Designing usable text*, (pp.245-312). Orlando: Academic Press.

Vaughn, W. (1992). Automated picture referencing: A further look at Morelli. *Computers and the History of Art, 2*(2), 7-18.

Wartofsky, M.W. (1979). Picturing and representing. In C.F. Nodine & D.F. Fisher (Eds.), *Perception and spictorial representation* (pp. 272-283). New York: Praeger.

Wolfe, T. (1981). *From Bauhaus to our house*. New York: Simon & Schuster.

ISSN: 2324-9773

Beebe, C., & Jacob, E. (1998). Graphic language documents: Structures and functions. 9th ASIS SIG/CR Classification Research Workshop, 29-52. doi:10.7152/acro.v9i1.12743

ISSN: 2324-9773